



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Competing evolutionary paths in growing populations with applications to multidrug resistance

Citation for published version:

Nicholson, M & Antal, T 2019, 'Competing evolutionary paths in growing populations with applications to multidrug resistance' PLoS Computational Biology. DOI: 10.1371/journal.pcbi.1006866

Digital Object Identifier (DOI):

[10.1371/journal.pcbi.1006866](https://doi.org/10.1371/journal.pcbi.1006866)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

PLoS Computational Biology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



RESEARCH ARTICLE

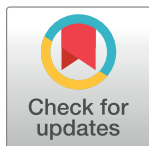
Competing evolutionary paths in growing populations with applications to multidrug resistance

Michael D. Nicholson^{1‡*}, Tibor Antal²

1 School of Physics and Astronomy, University of Edinburgh, Edinburgh, Scotland, **2** School of Mathematics, University of Edinburgh, Edinburgh, Scotland

‡ Current address: Department of Data Sciences, Dana-Farber Cancer Institute, Boston, Massachusetts, United States of America

* mdnicholson5@gmail.com



Abstract

Investigating the emergence of a particular cell type is a recurring theme in models of growing cellular populations. The evolution of resistance to therapy is a classic example. Common questions are: when does the cell type first occur, and via which sequence of steps is it most likely to emerge? For growing populations, these questions can be formulated in a general framework of branching processes spreading through a graph from a root to a target vertex. Cells have a particular fitness value on each vertex and can transition along edges at specific rates. Vertices represent cell states, say genotypes or physical locations, while possible transitions are acquiring a mutation or cell migration. We focus on the setting where cells at the root vertex have the highest fitness and transition rates are small. Simple formulas are derived for the time to reach the target vertex and for the probability that it is reached along a given path in the graph. We demonstrate our results on several scenarios relevant to the emergence of drug resistance, including: the orderings of resistance-conferring mutations in bacteria and the impact of imperfect drug penetration in cancer.

OPEN ACCESS

Citation: Nicholson MD, Antal T (2019) Competing evolutionary paths in growing populations with applications to multidrug resistance. PLoS Comput Biol 15(4): e1006866. <https://doi.org/10.1371/journal.pcbi.1006866>

Editor: Joachim Krug, University of Cologne, GERMANY

Received: October 8, 2018

Accepted: February 13, 2019

Published: April 15, 2019

Copyright: © 2019 Nicholson, Antal. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: MDN was funded by an EPSRC studentship, specifically research council award number 1527143. TA received no specific funding for this work. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

How long does it take for a treatment naive, growing bacterial colony to be able to survive exposure to a cocktail of antibiotics? En route to multidrug resistance, what order did the drugs become impotent in? Questions such as these that pertain to the emergence of a significant cell type in a growing population arise frequently. They are often investigated via mathematical modelling but biologically insightful results are challenging to obtain. Here we outline a general framework of a stochastically growing population spreading through a graph to study such questions and provide simple formulas as answers. The significant cell type appears upon the population reaching a target vertex. Due to their simplicity, the derived formulas are widely accessible and can be used to guide and develop intuition on a range of biological scenarios. We demonstrate this on several settings including: how a

region where drugs cannot penetrate affects the emergence of resistance, and, the ordering of mutations that leads to drugs being ineffective.

Introduction

The timing and manner in which a particular phenotype arises in a population is a central question of theoretical biology [1–13]. A typical scenario is to consider an initially monomorphic, wild type population, composed of cells that can acquire mutations, for example single site substitutions on the genome. The phenotype of interest comes to exist after a cell has accrued a specific set of mutations. The interpretations of this event are application dependent, but examples are the genesis of cancer instigated by mutations in a pair of tumour suppressor genes, or the emergence of multidrug resistance via alterations to the genes coding for the target proteins. Regardless of context, the questions of when, and how, the phenotype emerges are of significant interest.

It is commonly assumed that the population under consideration is of fixed size. However, with the aim of characterising disease progression, an increasing body of research has been developed to examine the evolutionary dynamics of a growing population. These studies have provided insights on a range of applications, including; cancer genetics [14–17], metastasis formation [18–22], drug resistance [23–26], phylogenetics [27], and the impact of poor drug penetration [28–31].

Here we continue in the same vein by considering a stochastically growing cellular population, where cells can transition in such a fashion so as to alter their, and their offsprings, reproductive capabilities. Such a transition might be due to the acquisition of a (epi)genetic alteration or migration into a new environment. As before, suppose we have a cellular state of interest, for example; a given genotype, a spatial location, or a combination of both. Will this state ever be reached? If it is, when is it reached? And by which sequence of intermediate states?

To make the discussion clear, let us consider an example application: the emergence of multidrug resistant bacteria. Suppose an infection begins with a single pathogenic bacterium which is sensitive to two antibiotic therapies, drug *A* and drug *B*. In the absence of either drug, the initial bacterium initiates a growing colony. Within the colony, when a cell replicates one of the daughter cells can acquire a mutation yielding resistance to either of the drugs. Here our questions are: how long does it take for multidrug resistance to emerge? En route, is resistance to drug *A* or drug *B* more likely to arise first? An ability to answer such questions is key to understanding pre-existing resistance, a common cause of therapy failure in various settings [32]. This scenario is illustrated in Fig 1a. There each vertex represents a cellular type, in this case its resistance profile. The edges represent cell transitions via mutation upon replication.

In this article we focus on the setting where the intermediate states have reduced reproductive ability (fitness) relative to the initial cells. This is primarily motivated by the commonly observed cost of resistance [33–35], whereby cellular populations that are resistant to a given drug grow more slowly than their sensitive counterparts. A second scenario where we expect a reduction in fitness is drug sensitive cells becoming exposed to toxins, which increases the cell's rate of dying (cytotoxic) or decreases the rate of replication (cytostatic). As in many biological applications the relevant transition rates are small (some representative examples are: measurements for the point mutation rate per cell division of 10^{-9} in cancer [36] and 10^{-10} for bacteria [37], while the dissemination rate of pancreatic cancer cells from the primary tumour was estimated as 10^{-7} per cell division in [19]), we concentrate our efforts on the regime of low transition rates. Our main contribution is to provide simple, intuitive formulas that answer the

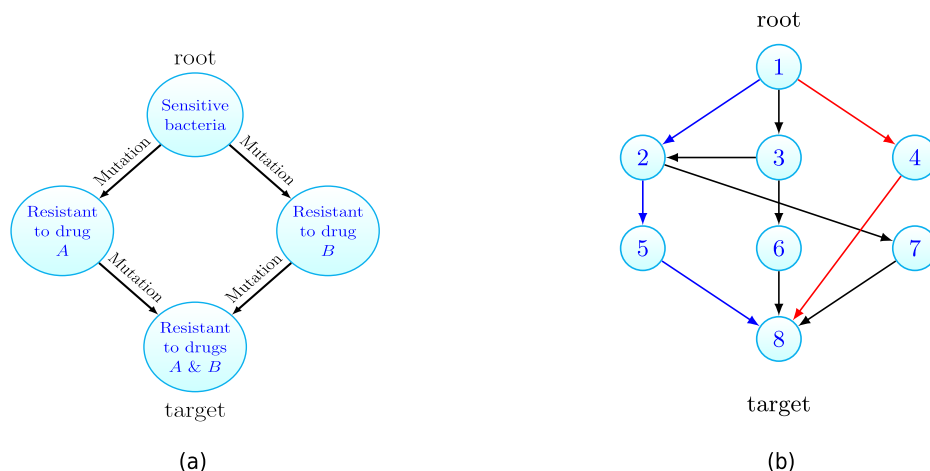


Fig 1. (a) Motivating example: consider a population of growing drug sensitive bacteria. Each vertex represents a cell type and cells can mutate to acquire resistance. Starting from one sensitive bacterium that can replicate, die or mutate, how long does it take to traverse the graph until multidrug resistance emerges? Is resistance to drug *A* then *B* more likely than the converse? (b) For the general case, the growing population is composed of different cell types that can be associated to vertices on a graph. Cells at vertex *x* can transition to vertex *y* if the edge (*x*, *y*) exists. Starting with cells at the root vertex, how long until the target is populated? Is the blue or red path more likely to initiate the target population? Here *N* = 8.

<https://doi.org/10.1371/journal.pcbi.1006866.g001>

questions posed in the small transition rate limit. These formulas show explicitly the contribution of the model parameters, e.g. transition rates and fitness reductions, which allow them to be probed for biological insight. This provides relationships which would be difficult to deduce from simulation alone.

We now move to detail the general framework we study, which will be seen to encompass models introduced in previous works as special cases [25, 26, 30, 38]. Our main results concerning when and how a particular cell type emerges are then presented, and we follow this by demonstrating our method on several applications.

Results

General framework

Our interest will always be in the emergence of a particular cell type in a growing population which is modelled as a specific form of a multitype branching process [39]. It is convenient to picture the population on a finite, simple, directed graph $G = (V, E)$, see Fig 1b, where each vertex of the graph represents a cell type or state. The number of types in the system is denoted *N* and so we let $V = \{1, 2, \dots, N\}$. Thus *E* is a subset of the set of ordered pairs $\{(i, j): i, j \in V, i \neq j\}$. For any given cell type there is an associated vertex in *V*. Henceforth we will refer to cells of type *x* as cells residing at vertex *x*. We will be concerned with the timing, and fashion, that vertex *N*, which we refer to as the target, is populated, when we initiate with cells at vertex 1, which we denote the root. In the example of multidrug resistance, illustrated in Fig 1a, the cells at the root are sensitive bacteria while the target population is resistant to both drugs.

Let (*x*) represent a cell at vertex *x* and \emptyset symbolise a dead cell. Then, with all cells behaving independently, our cell level dynamics can be graphically represented as:

$$(x) \rightarrow \begin{cases} (x), (x) & \text{at rate } \alpha(x) \\ \emptyset & \text{at rate } \beta(x) \\ (x), (y) & \text{at rate } v(x, y) \text{ if } (x, y) \in E. \end{cases}$$

That is: cells divide at rate $\alpha(x)$, die at rate $\beta(x)$, and transition to a cell at vertex y at rate $\nu(x, y)$ if the edge (x, y) exists. We will denote the fitness (growth rate) of cells at vertex x by $\lambda(x) = \alpha(x) - \beta(x)$, the difference of division and death rates of vertex x cells. The parameters associated with the vertex 1 population feature prominently and so for convenience we let $\alpha = \alpha(1)$, $\beta = \beta(1)$ and $\lambda = \lambda(1)$. As mentioned above, in this article we focus on the setting where the vertex 1 population has the largest fitness, which is also positive. That is, we assume that $\lambda > 0$ and for $2 \leq x \leq N - 1$, $\lambda(x) < \lambda$. We do not specify the fitness of the target population (cells at vertex N).

A common variant when modeling transitions (or mutations) is to specify that cells of type x divide at rate $\alpha'(x)$, and then with probability $\nu'(x, y)$ a transition occurs to vertex y . For a fixed value of parameters, this formulation of transitions is equivalent to that given above upon letting $\nu(x, y) = \alpha'(x)\nu'(x, y)$, $\alpha(x) = \alpha'(x)(1 - \sum_{y:(x,y) \in E} \nu'(x, y))$. Also in the model as currently stated, transitioned cells have deterministically fixed new division and death rates. However including a finite distribution of fitness effects, when a transitioned cell is assigned random division and death rates, is actually covered by our model as given thus far. We discuss this point further in the Extensions section.

At a population level, the number of cells at vertex x at time t will be denoted $Z_x(t)$. We shall always assume that the system is initiated with only vertex 1 cells, at a quantity z , that is $Z_x(0) = z\delta_{x,1}$ with the Kronecker delta function $\delta_{x,y}$.

We have two primary questions. The first is, having initiated the system with z vertex 1 cells, how long does it take for the population at the target vertex to arise? That is we concern ourselves with the distribution of the target hitting time, defined as

$$T = \inf \{t \geq 0 : Z_N(t) > 0\}. \quad (1)$$

Now assuming that the target vertex is populated by a founding cell, we ask from which path through the graph G did this founding cell come? This gives rise to a distribution over the set of paths (or walks) from the root vertex to the target vertex, which we aim to characterise. This second question is more precisely formulated in the Path distribution section. Throughout we will assume that we wait for the first cell at vertex N to arise, however it may be that the population founded by that initial cell goes extinct. If instead one wishes to wait for the first ‘successful’ cell, that is the first cell at vertex N to exist whose progeny survives, then all the results presented below hold so long as the mapping $\nu(x, N) \mapsto \nu(x, N)\lambda(N)/\alpha(N)$ for all target adjacent vertices x , is applied.

Our results are most clearly understood when G is acyclic, which amounts to excluding reverse transitions. Therefore in the following presentation this will be assumed. The setting when this assumption is relaxed will be discussed in the Extensions section.

Consider any path from the root to the target vertex. Intuitively, if the transition rates encountered along this path are larger, the target population will be seeded from this path quicker. Conversely, the smaller the growth rates along this path are, the slower the target will be reached. It will transpire that this intuition is indeed correct. The key quantities that show how the time to the target, and which path is taken, depend on these competing factors will be seen to be the path weights, which we now introduce.

Let the set of paths between the root and the target be denoted $\mathcal{P}_{1,N}$. Any $p \in \mathcal{P}_{1,N}$, will be a sequence $p = (p_1, p_2, \dots, p_l, p_{l+1})$, where each $1 \leq p_i \leq N$, l is the path length (the number of edges in p), all p_i are distinct (as presently G is acyclic), and $p_1 = 1$, $p_{l+1} = N$. Along the path p , let us call the difference between the growth rate at the root and the vertices encountered the

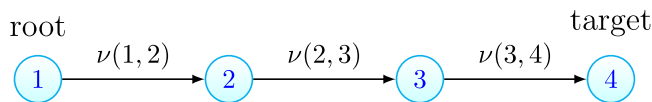


Fig 2. A path graph of length 3, with $N = 4$.

<https://doi.org/10.1371/journal.pcbi.1006866.g002>

fitness costs. Then we define the weight of path p as

$$w(p) = v(p_1, p_2) \prod_{i=2}^l \frac{v(p_i, p_{i+1})}{\lambda - \lambda(p_i)}, \quad (2)$$

that is the product of the transition rates along the path divided by the fitness costs along the path. Throughout the empty product is set to 1. Further, we let the total weight of the target be

$$\phi_N = \sum_{p \in \mathcal{P}_{1,N}} w(p). \quad (3)$$

We will take the case of a path graph as a running example. A path graph, as illustrated in Fig 2, is the scenario in which the only edges are $(i, i + 1)$, for $1 \leq i \leq N - 1$. There is only one path p between the root and the target and so in this case, $\phi_N = w(p) = v(1, 2) \prod_{i=2}^{N-1} \frac{v(i, i+1)}{\lambda - \lambda(i)}$.

Time until target vertex is populated

How long do we wait until the target vertex is populated? We answer this question by characterising the distribution of T , defined in (1), when the target seeding transition rates (the transition rates associated with edges into the target vertex) are small. The key tool required is the long-time population numbers at the initial and intermediate vertices, which is discussed further in the Materials and Methods section. Using this we are able to prove the following theorem regarding the target hitting time, whose proof can be found in S1 Appendix (see Sections S3 and S4).

Theorem 1. *As the target seeding transition rates tend to 0, we have*

$$\mathbb{P}(T - \mu > t) \rightarrow \left(\frac{\lambda/\alpha}{1 + e^{\lambda t}} + \beta/\alpha \right)^z \quad (4)$$

where

$$\mu = \frac{1}{\lambda} \log \frac{\lambda^2}{\alpha \phi_N}.$$

Heuristically, the time until the target is populated is approximately μ plus some noise, where the distribution of the noise is given by the right hand side of (4). Note that this distribution depends only on parameters associated with vertex 1 cells, while all other parameters of the system are bundled into μ . For practical purposes the theorem yields the following approximation

$$\mathbb{P}(T > t) \approx \left(\frac{\lambda/\alpha}{1 + e^{\lambda t} \phi_N \alpha / \lambda^2} + \beta/\alpha \right)^z \quad (5)$$

This approximation will be valid if the target seeding transition rates are small. We do not provide estimates on the approximation error, however simulations (for example see Fig 3b) demonstrate this approximation holds even for only moderately small transition rates.

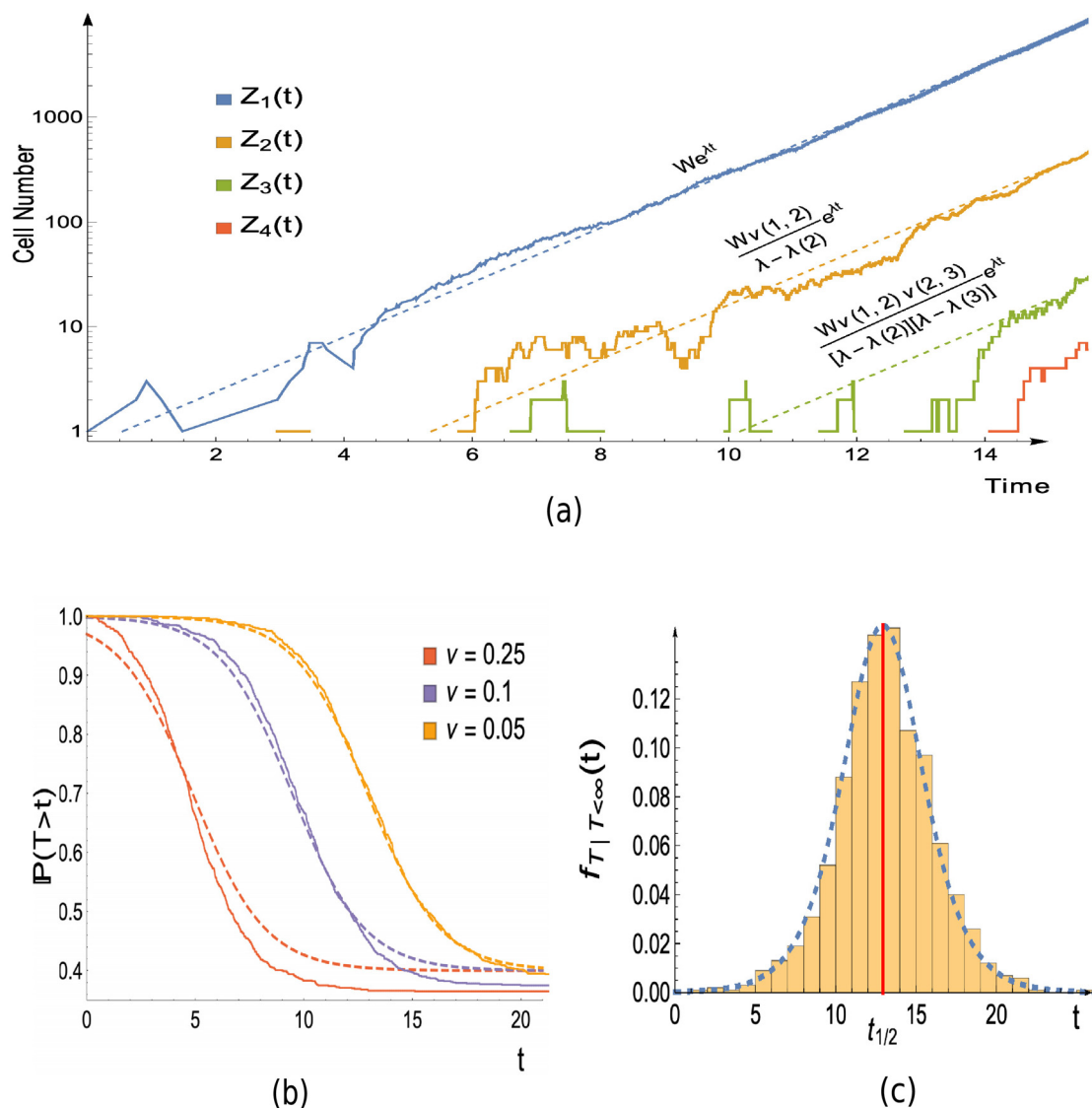


Fig 3. Simulation versus theory in the case of a path graph with 4 vertices, as in Fig 2. (a) Joined lines show a single realisation of the population growth obtained via simulation. Dashed lines are the asymptotic growth of the populations, which follows the functions given immediately above the dashed lines. These functions are a consequence of Theorem 3, which can be found in the Materials and Methods section. W is a random variable whose value will differ between simulations. Its value here is estimated by the realised value of $e^{-\lambda t_f} Z_1(t_f)$, with $t_f = 14.5$. (b) The distribution of the hitting time for vertex 4: Joined lines show the empirical distribution obtained from 1000 realisations, dashed lines display the theoretical approximation (5). The error in the approximation decreases as $v \rightarrow 0$, but still shows good agreement for moderate values of v . (c) Probability density function for hitting time of vertex 4, conditioned that it is finite: bars display normalised histogram constructed from 1000 realisations, while dashed line shows approximation which follows from differentiating the conditional variant of (5) (given in Eq. S79 of S1 Appendix). The approximate median time for the type 4 population to arise, as given (9), is indicated by the red vertical line. Parameters: $\alpha = \alpha(2) = \alpha(3) = \alpha(4) = 1$, $\beta = 0.4$, $\beta(2) = \beta(3) = 1.3$, $\beta(4) = 0.6$, $z = 1$; for (a) and (c), $v(1, 2) = v(2, 3) = v(3, 4) = 0.05$; for (b), legend displays value of $v = v(1, 2) = v(2, 3) = v(3, 4)$ used.

<https://doi.org/10.1371/journal.pcbi.1006866.g003>

Notice that it is possible that the target is never populated ($T = \infty$). In S1 Appendix (Proposition S1) we show that if the vertex 1 population survives forever then the target hitting time is finite with probability one. Furthermore in many relevant cases, namely a large initial population, low death rate, or small transition rates leaving vertex 1, that the target is eventually populated is essentially equivalent to the vertex 1 population surviving. The probability that

the vertex 1 population survives is obtained [S1 Appendix](#) (Eq. S7), and this quantity thus gives an approximation for the probability that the target population will ever arise, which is

$$\mathbb{P}(T < \infty) \approx 1 - (\beta/\alpha)^z. \quad (6)$$

This also explains the β/α term in (5), which arises due to lineages stemming from the initial vertex 1 cells going extinct. The corresponding approximate distribution for the target hitting time, when we condition that the vertex 1 population survives is given in [S1 Appendix](#) (Eq. S78). Differentiating the conditional hitting time distribution yields the density which is compared with simulations in [Fig 3c](#).

Let us now discuss the time centring term μ , as given in Theorem 1. In the case of $\beta = 0$, $z = 1$, it can be easily seen (recall $\lambda = \alpha - \beta$) that the median of our approximate distribution (5), is exactly μ . More generally, if we let $t_{1/2}$ be the median time to hit the target, then if the vertex 1 population survives, we have

$$t_{1/2} \sim \mu - h(z). \quad (7)$$

as the target seeding transition rates tend to 0. The shift $h(z)$, whose precise form can be found in [S1 Appendix](#) (Corollary S2), has the behaviour

$$h(1) = 0, \quad h(z) \sim \frac{1}{\lambda} \log \frac{z\lambda}{\alpha}, \quad z \rightarrow \infty. \quad (8)$$

The shift exists as initiating the system with a larger number of cells at vertex 1, leads to the target being reached faster. In terms of notation $f \sim g$ means that $f/g \rightarrow 1$, when the limit under consideration is taken.

We now return to the running example of the path graph setting which was introduced at the end of the General framework section. With $z = 1$, (7) yields that the median time for the target population to appear is

$$t_{1/2} \approx \mu = \frac{1}{\lambda} \log \frac{\lambda^2}{\alpha v(1, 2)} + \frac{1}{\lambda} \sum_{i=2}^{N-1} \log \frac{\lambda - \lambda(i)}{v(i, i+1)}, \quad (9)$$

for small $v(N-1, N)$. The first summand on the right hand side comes from waiting for the first transition from the vertex 1 population, while the second is due transitions between the remaining vertices. The first summand is distinct as the vertex 1 growth is the main cause of stochasticity, as discussed in the Materials and Methods section. The population growth and target hitting time is illustrated for the path graph case in [Fig 3](#).

Path distribution

We now move to our second question: which path leads to the target population arising? Our naive expectation is that paths with larger weights will be more likely. This simple conjecture turns out to be true. To show this we first introduce some notation.

To any existing cell, say at vertex x , we may define the cell's vertex lineage. This tracks the vertices of cells in the ancestral lineage of the cell under consideration and is a sequence of the form $(1, \dots, x)$. For example a cell with vertex lineage $(1, 3, 4)$ is at vertex 4, and descended from a vertex 3 cell, who in turn descended from a vertex 1 cell. Any vertex lineage is a path in G and we denote the number of cells with vertex lineage q at time t as $X(q, t)$. Note that while there may be multiple ancestors at the same vertex in a given cell's ancestral lineage, for example several generations of vertex 3 cells before the first vertex 4 cell, we do not record these in the vertex lineage (more precisely $q_i \neq q_{i+1}$).

Now for any root to target path $p \in \mathcal{P}_{1,N}$, that is a path of the form $p = (1, p_2, p_3, \dots, N)$, the first time that p is traversed can be defined as

$$T(p) = \inf\{t \geq 0 : X(p, t) > 0\}.$$

Observe that path p populating the target first is equivalent to $T(p) = T$. The question of which path initiated the target only makes sense if the target is eventually populated. To ensure that this occurs we condition on a surviving vertex 1 population, and hence let

$$\mathbb{P}_1(\cdot) = \mathbb{P}(\cdot | \text{vertex 1 population survives}).$$

Recall that the vertex 1 population surviving is used as a proxy for T being finite (the target is eventually populated), as discussed in the paragraph following (5). We can now state the answer to our second question, which is a special case of Theorem 2 presented below. The probability that the target is populated via path p is simply the path weight of p , suitably normalised, that is

$$\mathbb{P}_1(T(p) = T) \approx \frac{w(p)}{\sum_{q \in \mathcal{P}_{1,N}} w(q)} = \frac{w(p)}{\phi_N} \quad (10)$$

for small target seeding transition rates. One may ask not only if a particular path populated the target, but whether the target was initiated from a given set of paths, for an example see the Imperfect drug penetration: combination therapy section. In this case the probability that the target is populated via a particular set of paths is given by summing the normalised path weights over each path in the set.

If the vertex 1 population avoids extinction then $T(p)$ will be finite for each root to target path p . Therefore instead of only asking which path populated the target first, one might be interested in whether a particular path is traversed significantly faster. For example, does multidrug resistance obtained via one of the paths in Fig 1a occur t days before resistance from the other path. In order to also consider this case, for a root to target path p , we define

$$T(\neg p) = \min\{T(q) : q \in \mathcal{P}_{1,N} \setminus \{p\}\},$$

that is the first time the target is reached along any path except p .

We can now quantify the probability of reaching the target via path p more than t time units before any other path. To avoid discussing a technical assumption, which excludes pathological scenarios, the theorem is stated in approximate form. The more precise version can be found in S1 Appendix (Theorem S3 in Section S5).

Theorem 2. For small target seeding transition rates, and $t \in \mathbb{R}$,

$$\mathbb{P}_1(T(\neg p) - T(p) > t) \approx \frac{w(p)}{w(p) + e^{\lambda t} \sum_{q \in \mathcal{P}_{1,N} \setminus \{p\}} w(q)}.$$

Note that letting $t = 0$ gives (10). In the running example of a path graph we cannot ask which path seeded the target vertex. However we briefly illustrate the usefulness of the above theorem by considering the question of whether the target vertex is populated via a fitness valley (where by fitness valley we mean a path such that cells at the internal vertices have smaller growth rates than the root population).

Let us consider the case depicted in Fig 4a where two paths exist to the target, a direct path and an indirect path. Here $N = 3$, with paths $p^{(1)} = (1, 2, 3)$, $p^{(2)} = (1, 3)$ and $v = v(1, 2) = v(2, 3)$, $v(1, 3) = v^2$. Naively it may be expected that as the product of transition rates along both these paths are equal, and the indirect path ($p^{(1)}$) contains a vertex with a fitness cost, the direct

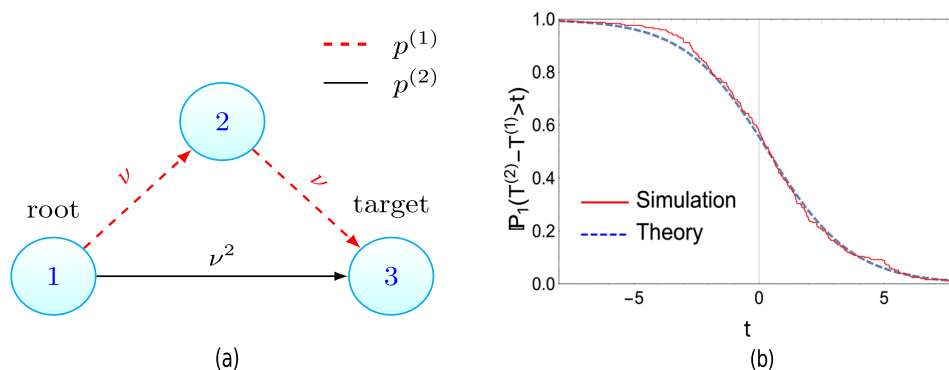


Fig 4. (a) Crossing valleys can be faster. Despite the red-dashed path containing a fitness valley (cells at vertex 2 have a reduced growth rate compared to the root population), the target population is more likely to arise via this path if $\lambda - \lambda(2) < 1$, as derived in (11). This is confirmed by the simulations displayed in (b), where $\lambda - \lambda(2) = 0.8 < 1$ and the Monte-Carlo estimate for $\mathbb{P}_1(T^{(2)} > T^{(1)})$ (the probability path 1 seeds the target first) is 0.58, while the analytic result yields 0.56. (b) Stochastic simulations of system in (a) vs theory (11). Parameters $\alpha = 0.9, \beta = 0.3, \lambda = 0.6, \alpha(2) = 0.2, \beta(2) = 0.4, \lambda(2) = -0.2, v = 0.1$, runs = 250.

<https://doi.org/10.1371/journal.pcbi.1006866.g004>

path ($p^{(2)}$) is more likely. However, for small v , Theorem 2 informs us that the target is populated via the indirect path t time units before the direct path, with probability

$$\mathbb{P}_1(T^{(2)} - T^{(1)} > t) \approx [1 + e^{\lambda t}(\lambda - \lambda(2))]^{-1} \quad (11)$$

where $T^{(i)} = T(p^{(i)})$. Thus the indirect path is more probable, that is $\mathbb{P}_1(T^{(1)} < T^{(2)}) > \mathbb{P}_1(T^{(2)} < T^{(1)})$, if $\lambda - 1 < \lambda(2)$.

Further applications of the results presented thus far will be given in the Applications section. Before this, we discuss the case when G is cyclic, and some extensions to the initial model for which our results still hold.

Extensions

Cyclic graphs. In the above exposition we have assumed that G is acyclic, that is no path exists which contains cycles. We now briefly summarise how the results are altered if G is cyclic, with details given in S1 Appendix. We will however maintain that the target is a sink vertex (has no outgoing edges) and the root is a source vertex (has no incoming edges). The condition on the root is perhaps unnatural, as we now allow back transitions between vertices except to the root, but is required for technical reasons which are discussed in the S1 Appendix (see the remarks following the proofs of Theorem S1 and Proposition S2). We aim to drop this condition in future work.

First we introduce the $N - 1 \times N - 1$ matrix A with entries

$$a_{ij} = \begin{cases} \lambda(i) & j = i \\ v(j, i) & \text{if } (j, i) \in E \\ 0 & \text{otherwise} \end{cases}$$

and we denote the largest real eigenvalue of A as λ^* .

In order that a modified version of Theorem 1 holds it is required that $\lambda^* = \lambda$ and λ^* is a simple eigenvalue. This condition is guaranteed for acyclic graphs, and will also be satisfied if the transition rates throughout the graph are small enough (see Lemma S2 in the S1 Appendix). Assuming this eigenvalue condition, Theorem 1 holds but with ϕ_N replaced by summing particular elements of the λ -associated eigenvector of A (see Theorem S2 in S1 Appendix).

Theorem 3, which concerns the population growth at the initial and intermediate vertices and is presented below in the Materials and Methods section, is similarly modified under the same condition (see Theorem S1 in S1 Appendix).

Turning to Theorem 2, we can now consider walks that contain cycles as opposed to only paths in the graph G . If we consider a particular set of root to target walks of a finite length, analogously to Theorem 2, we can ask whether the target is reached via one particular walk more than t units before any other walk in the specified set. A near identical result to Theorem 2 is obtained (see Proposition S2 in S1 Appendix). However the path weights should now be replaced by the walk weights (defined analogously to the path weight; the product of the transition rates along the walk divided by the fitness costs along the walk). In particular for two given walks of finite length we can still approximate the probability the target population arises via one of the walks versus the other.

For example we can use this to explore when a walk, containing a cycle, will seed the target before a path. Consider the graph and edge parameters displayed in Fig 5a. Let $p = (1, 2, 4)$ and $\omega = (1, 2, 3, 2, 4)$. Despite the extra length, by using the modified Theorem 2, for small v_1 we see

$$\mathbb{P}_1(T(\omega) < T(p)) \approx \left(1 + \frac{(\lambda - \lambda(2))(\lambda - \lambda(3))}{v_2^2}\right)^{-1}. \quad (12)$$

Hence the target can be seeded via the walk, if v_2 is large enough. On the other hand, by using the modified Theorem 2 it can be seen that when all the transition rates are small in the system, the target population will arise from a path and not a walk containing cycles (see Proposition S3 in S1 Appendix). This further justifies our focus on the acyclic case.

Distribution of fitness effects. Let us now expand on the comments in the General framework section where we claimed that having a finite distribution of fitness effects is covered by our model. This may be of interest, for example, when a cell acquires drug resistance via a mutation, but the cost of resistance is stochastic. More generally, suppose we wish to consider cells transitioning from vertex x to vertex y but require that the new growth rate of cells to be $\lambda_i(y)$ with probability $\pi_y(i)$, where $\pi_y(i)$ is a finite distribution and every $\lambda_i(y) < \lambda$. Then we can simply replace the vertex y with new vertices associated to each of the $\lambda_i(y)$ and have cells transition from vertex x to the new vertices at rate $v(x, y)\pi_y(i)$. We may freely apply the

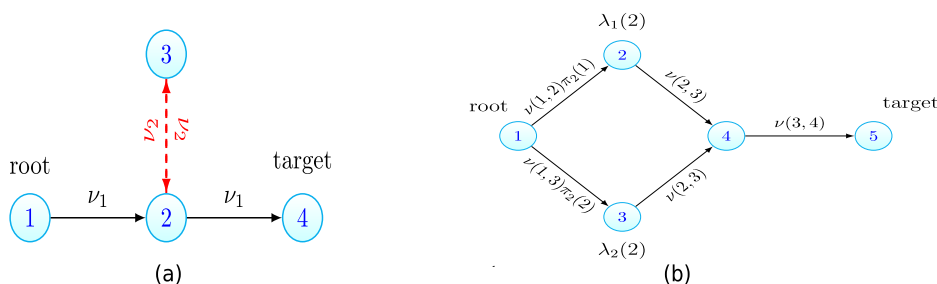


Fig 5. Example systems for which our extended results, given in the Extensions section, are useful. (a) Walks containing cycles can seed the target. The walk of length 4 that traverses the red edges with transition rate v_2 can still be faster than the length 2 path. See (12). (b) Incorporating a distribution of fitnesses. Take the path graph from Fig 2. Suppose that instead of a fixed $\lambda(2)$, after the first transition cells have fitness $\lambda_1(2)$ or $\lambda_2(2)$ with probabilities $\pi_2(1)$, $\pi_2(2)$, with $\lambda_i(2) < \lambda$. Then instead of the original path graph, we consider the enlarged graph shown here, where vertex 2 is replaced by two vertices with associated growth rates $\lambda_1(2)$, $\lambda_2(2)$. The transition rates to these new vertices are $v(1, 2)\pi_2(i)$, where $v(1, 2)$ is the transition rate between vertices 1 and 2 in the original setting. All results can then be applied to the new enlarged graph.

<https://doi.org/10.1371/journal.pcbi.1006866.g005>

results of the Results section to the graph containing the new vertices and edges. See Fig 5b for an example.

Alternative offspring distribution. We have focused on the setting where cells can divide, die or transition, however Theorem 2 would also hold for a slightly more general branching process. Keeping the transition process as it is, we can have a vertex x cell being replaced by j vertex x cells at a rate $\alpha(x)\sigma_{x,j}$ where for each x between 1 and $N - 1$, $(\sigma_{x,j})_{j \geq 0}$ is a probability distribution with finite second moment and mean $\bar{\sigma}_x$. We now have no need for $\beta(x)$ as cell death is contained in $\sigma_{x,0}$. This scenario may be relevant to the case of viral dynamics, where an infected cell can release multiple virions upon bursting [40]. In this setting if we redefine $\lambda(x) = \alpha(x)(\bar{\sigma}_x - 1)$ then the path distribution will be unchanged. This is due to the fact that the key result of [41] used in the proof of Theorem 3 holds then also. However Theorem 1 is no longer true in this setting. This is due to a lack of understanding of the large time behaviour of the root population (see the remarks following the proof of Theorem S1 in S1 Appendix).

Discussion

Before discussing related work and summarising this study, we demonstrate how to apply the results of our Results section on some applications. The approximate formulas given in the Results section will be our key tools, and we now briefly remark on these.

The results of the Results section hold when the final transition rates in each path tends to 0. Therefore in our applications all transition rates, in particular the mutation and migration rates discussed below, will be taken to be small and statements are to be interpreted as approximations that hold true in this limiting regime. We also note that in our model, when transitions represent migrations, they also occur at cell divisions, $(x) \rightarrow (x)$, (y) , instead of at any time, as $(x) \rightarrow (y)$. This might appear strange from a biological point of view, but the former formulation for transitions has been chosen, as it simplifies the mathematical treatment. More importantly, we expect that for small transition rates these formulations lead to very similar target hitting times. Indeed, simulations support this claim, as presented in Fig B of S1 Appendix. In all applications considered we shall neglect the role of back transitions. This is for simplicity and is in keeping with the previous works that we compare with. The effect of (finitely many) back transitions could be included by using our results on cyclic graphs.

Applications

First we consider the impact of imperfect drug penetration on the emergence of resistance. Two recent publications have explored how resistance spreads in this setting, and have shown that poor penetration can accelerate resistance [30, 31]. We are able to recover and extend some of their findings in a rigorous fashion. Next, the ordering by which resistance-causing mutations accrue is investigated, firstly in the setting of cancer and then in bacterial infections. In the case of bacterial infections we examine how the risk of multidrug resistance depends on mutation rates, recovering simulation results in [26]. A particular aim of this section is to illustrate how to apply the results of our Results section, and so, for the readers' convenience, we have collected the key formulas in Table 1.

Imperfect drug penetration: Monotherapy. Our first two applications concern imperfect drug penetration and so the language introduced here will hold in the next section also. The first scenario we examine is exactly the same as that considered in [30].

We consider a pathogenic population, with cancer cells or bacteria as example populations, being treated with one or more drugs with imperfect penetration profiles. The imperfect penetration results in low drug concentration in spatial locations, such that cells in these locations

Table 1. Key formulas needed for the Applications section. These approximations will be valid for small transition rates. The approximation for $t_{1/2}$ assumes a large initial number of cells at the root, $z \gg 1$. For small z see Corollary S2 in S1 Appendix.

Description	Formula	Reference
Weight of path p	$w(p) = v(p_1, p_2) \prod_{i=2}^l \frac{v(p_i, p_{i+1})}{\lambda - \lambda(p_i)}$	Eq (2)
Total weight of target	$\phi_N = \sum_{p \in \mathcal{P}_{1,N}} w(p)$	Eq (3)
Distribution of target hitting time	$\mathbb{P}(T > t) \approx \left(\frac{\lambda/z}{1 + e^{d\phi_N z/\lambda^2}} + \beta/\alpha \right)^z$	Eq (5)
Median of target hitting time	$t_{1/2} \approx \frac{1}{\lambda} \log \frac{\lambda}{z\phi_N}$	Eq (7)
Probability target populated via path p	$\mathbb{P}_1(T(p) = T) \approx \frac{w(p)}{\phi_N}$	Eq (10)

<https://doi.org/10.1371/journal.pcbi.1006866.t001>

may still proliferate. We term the cumulation of these spatial regions the *sanctuary*. Sanctuaries have been observed in both bacterial infections [42, 43] and tumours [44]. We consider resistance to have arisen when cells with sufficiently many mutations such that all therapies are ineffective come to exist in areas where the drugs have penetrated. The possibility of resistance occurring prior to treatment is excluded and we consider only the case when resistance arises from the sanctuary. Further, we do not consider cells acquiring mutations via gene transfer.

Throughout this section and the next we shall be comparing paths composed of mutation and migration events. Growing cells in the sanctuary will be the vertex 1 cells from the General framework section and thus have division and death rates α, β , with growth rate λ . The lack of exposure to drug motivates $\lambda > 0$. We shall assume all cells have division rate α . Across all cell types, i.e. regardless of location or drug presence, the per cell mutation and migration rates will be denoted v and m respectively. We fix the initial number of cells in the sanctuary to be z . We will often suppose that z is dependent on how penetrative a particular drug is (as this will affect the sanctuary size).

In the case of monotherapy, there are two paths from which resistance might arise. Firstly, upon replication a cell in the sanctuary may produce a mutant able to grow in the presence of the drug. Then a mutated cell in the sanctuary can migrate to the drug compartment. We term this the mutation-migration path. Analogously we have the migration-mutation path which involves a susceptible cell migrating from the sanctuary, from whose progeny a resistant mutant emerges. We suppose a resistance-conferring mutation carries a fitness cost which will be denoted s . Resistance costs may not always exist but are well documented and often assumed in viruses, bacteria and cancer [33–35]. Due to the fitness cost of resistance the death rate of mutated cells in the sanctuary is increased to $\beta + s$, while the death rate of migrated susceptible cells in the drug compartment is $\beta + d$, with $s, d > 0$.

The first quantity to determine is the probability of escaping resistance, and its dependence on the initial number of cells in the sanctuary. The approximation of (6) shows that the probability resistance never occurs decreases exponentially with the initial number of cells in the sanctuary as $\mathbb{P}(\text{resistance never occurs}) \approx (\beta/\alpha)^z$. For the rest of the discussion we will assume that resistance does occur. Further we suppose $z \gg 1$, so that we may use the approximate form for $h(z)$ given in (8). The case when z is small could be treated similarly using the exact expression for $h(z)$.

We now ask when, and via which path, will resistant cells come to exist in the drug compartment? From the Results section we know the key quantities required to answer these questions are the path weights. For this system these are given by

$$w(\text{mutation-migration path}) = \frac{vm}{s}, \quad w(\text{migration-mutation path}) = \frac{vm}{d}. \quad (13)$$

If we now let $t_{1/2}$ be the median time resistance occurs, then using the path weights with (7), we immediately have

$$t_{1/2} \approx \frac{1}{\lambda} \log \frac{\lambda^2}{\alpha} - \frac{1}{\lambda} \log \left(\frac{vm}{s} + \frac{vm}{d} \right) + \frac{1}{\lambda} \log \frac{\alpha}{z\lambda} \quad (14)$$

Of note is that the sanctuary size, mutation and migration rate all have equivalent impact. This simple result can now be explored under different biological scenarios.

For example, suppose we are currently using drug A with fitness cost $d^{(A)}$ and penetration profile leaving $z^{(A)}$ initial cells in the sanctuary. We consider instead using a second drug which results in sanctuary size $z^{(B)}$ and has fitness cost $d^{(B)}$. Let the median resistance times under the first and second drug be denoted $t_{1/2}^{(A)}$ and $t_{1/2}^{(B)}$ respectively. Then, using (14) and examining $t_{1/2}^{(B)} > t_{1/2}^{(A)}$, we find we should switch to drug B, if

$$\frac{z^{(A)}}{z^{(B)}} > \frac{1 + s/d^{(B)}}{1 + s/d^{(A)}}. \quad (15)$$

This inequality is illustrated in Fig 6b.

Note as $d^{(B)} \rightarrow \infty$, implying drug B is far more effective, the condition tends to $z^{(A)}/z^{(B)} > (1 + s/d^{(A)})^{-1}$. This is due to the fact that when $d^{(B)}$ is large resistance will always arise via the mutation-migration path (which we show below), so this provides an upper bound to the gain that can be obtained from increasing the drug efficacy. The above is under the assumption resistance will occur but note the probability of no resistance would decrease exponentially with larger sanctuary size.

Having discussed the time resistance will occur we now ask by which of the two paths it will occur. Let T^{vm} be the first time resistance arises via the mutation-migration path and T^{mv} be defined analogously for the migration-mutation path. Then Theorem 2 gives us

$$\mathbb{P}(T^{vm} < T^{mv}) \approx \frac{d}{s + d}.$$

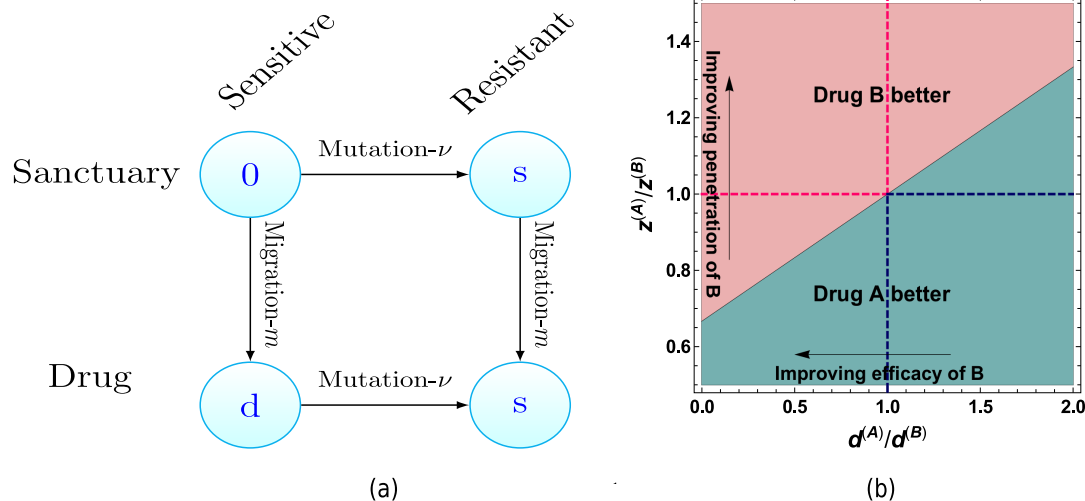


Fig 6. (a) In the case of monotherapy, there are two paths by which drug resistant cells can arise in the region containing the drug. Horizontal edges represent changes in genotype (gaining resistance) whereas vertical edges represent changes in spatial location. The labels of vertices are the fitness cost cells experience in that state and μ , m are the per cell mutation and migration rates. (b) Comparison of two drugs with differing efficacies and penetration profiles using (15). Arrows pertain to characteristics of drug B relative to drug A. Top left and bottom right quadrants illustrate region where drugs are trivially better, i.e. have superior penetration and efficacy. Parameters: $d^{(A)} = 10^{-2}$, $s = 5 \times 10^{-3}$.

<https://doi.org/10.1371/journal.pcbi.1006866.g006>

Hence resistance arising via the mutation-migration path is more probable, that is the above probability is greater than $1/2$, if $d > s$. As most drugs are designed to at least halt cell growth, while many resistant cell lines can still proliferate, we expect this condition to be satisfied for most drugs. Note that neither m nor v appear in the above expression as the numerator of both path weights is equal, see (13), and hence cancels.

A similar question was asked in [30]. There they considered an approximate stochastic process and sought the parameter regime such that T^{mv} stochastically dominated T^{mm} under the approximate process. For the full process considered here, we can ask the same question but on the limit of the centred (in the sense of Theorem 1) resistance times. The resulting condition for stochastic dominance is again $d > s$. The difference in the condition given here ($d > s$) and that stated in equation (28) in [30] exists as there v is defined as the probability of mutation per division event. Accounting for this leads to the same condition.

Imperfect drug penetration: Combination therapy. There are at least two reasons for extending our analysis to include two drugs, say drugs A and B, being used in tandem. Firstly, combination therapies are widely implemented in HIV and bacterial infections, and there is a growing appreciation for their use in cancer [23, 45, 46]. Secondly, interesting new phenomena emerges, especially when a region still exists with only one drug present. The work presented in this section is inspired by the model and questions of [31].

We will only consider the case where the penetration profile for drug B is a subset of drug A, and thus a single drug region exists only for drug A. This allows us to clearly demonstrate the impact of unequal penetration profiles. In this setting there are 12 possible states that a cell may inhabit, defined by which of the drugs the cell is exposed to and the resistance profile it has obtained. These states are illustrated in Fig 7. We further assume that the mutation rate bringing resistance to either drug is v . This is for simplicity only, and dealing with differing mutation rates is straightforward. Mutations that confer resistance will again have a fitness

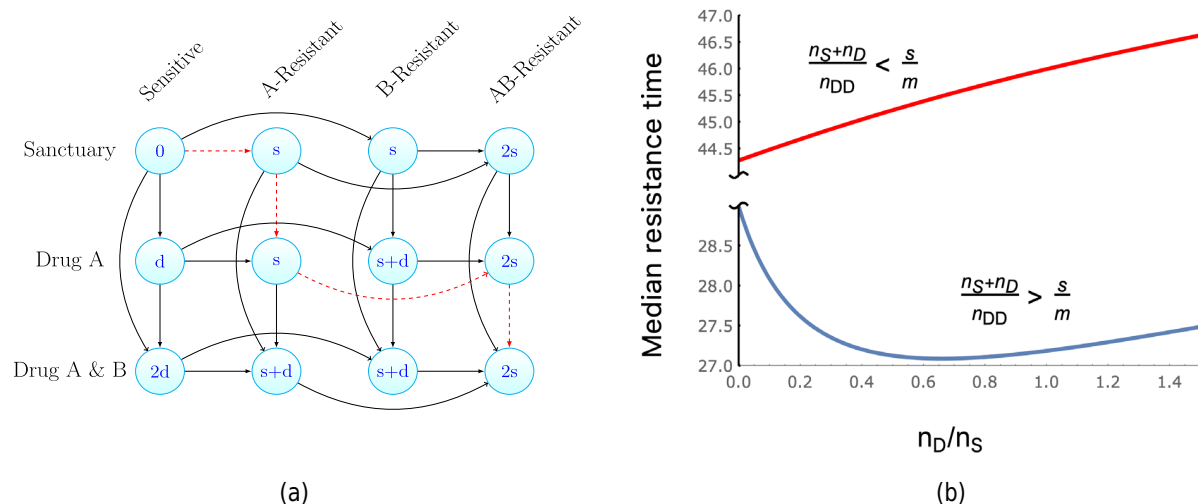


Fig 7. (a) Illustration of the paths to multidrug resistant cells arising in the region containing both drugs A and B, as in Fig 6. Note a spatial location exists containing only drug A. The red, dashed path is an example of stepwise evolution, where multidrug resistant cells come to exist in the sanctuary via the region containing only a single drug. The vertex labels display the fitness cost (increase to the cellular death rate) of cells in that state. (b) Exploration of the median time to resistance as the size of the drug A region increases relative to the sanctuary size. Here we plot $t_{1/2}$, as given in Table 1, for the system displayed in (a). We observe that, for a subset of the parameter space, the existence of a single drug region can decrease the time to resistance, despite diminishing the sanctuary size. If both drugs penetrate the majority of the system and the drugs have a greater effect than the cost of resistance, this subset of parameter space is identified by the inequality (16). Within this parameter subset, resistance is fastest at a particular size of single drug compartment, given by (17). Parameters; (top) $v = 10^{-6}$, $m = 0.05$, $d = 0.9$, $s = 10^{-3}$, $\lambda = 0.4$, $\alpha = 0.3$, $n_{Tot} = 107$, $n_{DD} = 0.9n_{Tot}$, $\gamma = 10^{-3}$; (bottom) all same except $s = 10^{-2}$.

<https://doi.org/10.1371/journal.pcbi.1006866.g007>

cost s , and sensitive cells coming into contact with drugs invokes a fitness cost d . These fitness costs increase the cellular death rate additively. The increase to the death rate for a cell in any given state is displayed as the vertex labels in Fig 7. Drug interactions, mutation specific fitness costs, and epistasis are neglected but may easily be included.

To investigate the effect of the single drug region, we introduce further notation. Let n_{Tot} be the number of cells that may reside in our system of interest (e.g. the entire patient or tumour) in the absence of drug. Further, let the regions where both drugs act, the single drug (drug A) acts and the sanctuary have their respective cell capacities denoted by n_{DD} , n_{D} , and n_{S} . We assume that all these capacities are large, and hence we may still investigate the system with our branching process model. These capacities enter the dynamics of the process in the following manner. When a cell leaves the region it is in, at rate m , it now may migrate to one of two regions. Of these two regions, we specify that it migrates to a particular region at a frequency proportional to the regions capacity. For example a cell in the sanctuary migrates to the double drug region at rate $mn_{\text{DD}}/(n_{\text{D}} + n_{\text{DD}})$. Resistance mutations and the presence of the drug has the same additive effect as in the Imperfect drug penetration: monotherapy section.

Observe that there are 18 possible paths by which sensitive cells in the sanctuary can eventually produce fully drug resistant cells in the double drug compartment. To see this it is convenient to separate the paths into those that pass through the single drug compartment, which following [31] we term stepwise evolution, and those that involve migration from the sanctuary directly to the double drug compartment, which we collectively call direct evolution. For the direct evolution paths there are three transitions needed, mutations yielding resistance to drug A and B and then the migration to the double drug region. From the number of permutations of these events, this yields 6 direct evolution paths. In the stepwise evolution case four transitions are needed, both mutations and then the two requisite migration events. Notice that the migration from the sanctuary to the single drug region must precede the migration from the single drug region to the double drug region. This leads to 12 stepwise evolution paths. An exemplar stepwise evolution path is a mutation to drug A, followed by migration to the single drug region, then a mutation to drug B, and finally a migration to the double drug region. The edges of this path are red and dashed in Fig 7a.

Suppose, momentarily, that the penetration profiles of drugs A and B exactly match (but are still imperfect). In this case, resistance in the double drug region can be obtained only via direct evolution. If we now increase the penetration of drug A only, 12 new possible paths to resistance open up (the stepwise evolution paths of the previous paragraph), accelerating the time until resistance. However, the increased penetration also leads to a reduction in the sanctuary size, which should delay resistance. Hence a trade-off may exist with respect to the penetration profile of drug A. To study this we fix $n_{\text{Tot}} = n_{\text{S}} + n_{\text{D}} + n_{\text{DD}}$ and specify that the initial number of cells in the sanctuary is related to its size via $z = \gamma n_{\text{S}}$ with $0 < \gamma \ll 1$. Further let $t_{1/2}(n_{\text{D}}/n_{\text{S}})$ be the median time to resistance with fixed $(n_{\text{D}}/n_{\text{S}})$. Using the approximation provided by (7) we can explore the behaviour of $t_{1/2}(n_{\text{D}}/n_{\text{S}})$. As can be seen in Fig 7b it is possible, in certain regimes of the parameter space, that the existence of the single drug compartment accelerates resistance (by reducing $t_{1/2}$). We now seek the conditions under which this is the case. Brief derivations of the presented formulas below (16, 17, 18) are given in S1 Appendix (Section S8).

As mentioned previously, it is reasonable to expect the efficacy of either drug is greater than the cost to resistance. Also we may hope that both drugs penetrate the majority of the target system. This motivates taking $s \ll d$ and $(n_{\text{S}} + n_{\text{D}}) \ll n_{\text{Tot}}$. Under this limit, by examining the sign of $\frac{d}{d(n_{\text{D}}/n_{\text{S}})} t_{1/2}(0)$, we find a simple condition for when increasing the single drug

compartment accelerates resistance, namely

$$\frac{n_s + n_D}{n_{\text{Tot}}} > \frac{s}{m}. \quad (16)$$

However, when (16) holds, as n_D/n_s increases, eventually a minimal resistance time is achieved. This represents a worse case scenario yielding fastest possible resistance. For $s \ll d$, this minimal time occurs when

$$\frac{n_D}{n_s} \approx 1 - \frac{2sn_{\text{Tot}}}{m(n_s + n_D)}. \quad (17)$$

We now examine whether resistance arises via stepwise evolution or direct evolution. Let T^{SE} be the first time resistance occurs via stepwise evolution, and T^{DE} defined analogously for direct evolution. Then using Theorem 2, again with $s \ll d$, we find

$$\mathbb{P}(T^{\text{SE}} < T^{\text{DE}}) \approx \left[1 + \frac{s(n_{\text{DD}} + n_s)}{mn_D} \right]^{-1}. \quad (18)$$

Therefore the larger the proportional size of the single drug region, the more likely resistance will arise via stepwise evolution. A qualitatively similar result was derived in [31] (e-page 2878) by comparing the most likely stepwise evolution path with the most likely directed evolution path.

Furthermore, in [31] a minimal resistance time, in the sense of (17), was observed via simulations for the mutation_A-migration-mutation_B-migration path (the red and dashed path in Fig 7a). Using Theorem 2 we can show that this is the most probable (in the sense of the mode of the path distribution, see (10)) stepwise evolution path whenever $d > s$. Further, by differentiating the median time along this path, we see a minimal time occurs at $n_D \approx n_s$, which recovers the simulation result in [31] (see their Fig. 3). That this path is the most probable stepwise evolution path also gives insight into the recurring s/m term in (16) and (18). Relative to direct evolution this path has an extra migration step, so higher migration should increase the probability of seeding the target via this path. However fitness costs along this paths are incurred only via the cost of resistance, and a smaller s diminishes these fitness costs. While the above is motivated by and compared with the study [31], in contrast to that work, we do not consider a size-dependent branching process (the population was constrained by a large carrying capacity in [31]). Despite this we are able to derive similar qualitative, and indeed quantitative (e.g. the recovery of the simulation result discussed above), results. These features may be general properties of such models or are perhaps due to the size of the carrying capacity in [31].

Path to resistance in chronic myeloid leukemia. Let us discuss the example application of [25], concerning the mechanism by which resistance to the targeted therapy imatinib occurs in chronic myeloid leukemia (CML). The two main resistance mechanisms are via gene amplification of the BCR-ABL fusion gene, which drives the CML, or a point mutation resulting in a modification to the target protein of imatinib. Approximately 100 point mutations have been identified conferring resistance, and a rough estimate for the probability of each of these occurring during a cell division is 10^{-7} [47]. Thus the probability of a point mutation conferring resistance is approximately 10^{-5} per division. The analogous quantity for resistance-causing gene amplifications, in a similar system, was found to be approximately 10^{-4} [48]. With the division rate of leukemic stem cells as $\alpha = 0.008 \text{ day}^{-1}$ [49] we have the rate of point mutations as $v_{pm} = 8 \times 10^{-8} \text{ day}^{-1}$, while the rate of resistance due to gene amplifications is $v_{ga} = 8 \times 10^{-7} \text{ day}^{-1}$. Therefore, using this information alone, the probability of resistance arising from a

point mutation first is estimated to be 1/11. However, in the majority of cases, the primary mechanism of resistance is found to be point mutations [50].

Within our framework there are two possible explanations we can explore, both related to the reproductive capabilities of resistant cells. The first is simply that more lineages of amplified cells go extinct. The paragraph directly following (1) indicates how to treat this case. Indeed, if ρ_{pm} is the average survival probability for a population initiated by a single cell with a point mutation and ρ_{ga} the analogous quantity for gene amplification, then we find resistance arising first from a point mutation is more probable if $\rho_{pm} > 10\rho_{ga}$. This is possible as point mutations may be deleterious or advantageous [47] whereas gene amplifications appear deleterious [51].

The second explanation, that suggested in [25], is that several gene amplifications are required to attain resistance. Suppose a gene amplification increases the death rate of cells by s per day. Then if two gene amplifications are required, ignoring the survival probability from before, the probability that resistance arises via point mutation is now

$$\frac{v_{pm}}{v_{pm} + v_{ga}^2/s} = (1 + 8 \times 10^{-6}/s)^{-1}.$$

Resistance via points mutations is now more likely if $s > 10^{-6}$. Note that Fig 4a serves as a schematic for this scenario, reinterpreting the fitness valley path as resistance arising via gene amplifications and the direct path relating to resistance via point mutations. Further explanations for the primacy of point mutation mediated resistance exist and the assumptions above may not hold. But when they do hold, our results offer a clear framework to investigate such questions.

Antibacterial multidrug resistance. A timely problem is that posed by the emergence of multidrug resistant bacteria. We now explore some specific examples of our general framework that pertain to this issue. In particular we consider the accumulation of resistance-conferring mutations as a bacterial colony grows in the absence of drugs. As before we assume resistance incurs a fitness cost and thus the wild-type (unmutated cells) has the highest growth rate. Using Theorem 2 and measured fitness costs and mutation rates we can predict the most likely path to multidrug resistance.

We consider the acquisition of resistance to both rifampicin and streptomycin in *P. aeruginosa*. Recently the fitness costs associated with resistance to rifampicin [52] and streptomycin [53] have been reported. Fitness costs were determined by the maximum growth rate of the bacteria undergoing exponential growth on nutrient-rich plates. From [53] the ratio of streptomycin resistant bacteria to the wild-type's growth rate was 0.71 (here the growth rate of the wild-type is estimated from Fig. 2.a in [53]). The same value for resistance to rifampicin, in [52], was 0.88. Let T^{RS} be the time when multidrug resistance emerges assuming rifampicin resistance is acquired first, and T^{SR} defined analogously for streptomycin resistance first. Then from Theorem 2, we have that

$$\mathbb{P}(T^{RS} < T^{SR}) \approx \frac{1 - 0.71}{(1 - 0.88) + (1 - 0.71)} = 0.7.$$

Hence, under this model, we would predict that rifampicin resistance emerges first en route to multidrug resistance. We note that above we have taken the average fitness costs associated with the differing antibiotics. Fluctuations have not been taken into account.

Instead of considering the the acquisition of resistance to different drugs within a bacterial strain, we can also examine the impact of differing mutation rates across strains. In [26] the heightened occurrence of multidrug resistance in lineage 2 vs lineage 4 from *M. tuberculosis*

was investigated. Considering in particular resistance to rifampicin and isoniazid, via simulations they deduced that the heightened occurrence of resistance was due to elevated mutation rates in lineage 2 strains. The model simulated can be considered a discrete time counterpart to ours in the case of two paths of length 2 (see Fig 1a). One of the main quantities of interest was the relative probability of resistance, when the only difference between the strains was the mutation rates. For a particular strain, initiated with one wild-type cell with wild-type division and growth rates, α, λ , suppose we have mutation rates v_R, v_I to rifampicin and isoniazid, each incurring the same fitness cost s . Then from Theorem 1, we have

$$\mathbb{P}(T \leq t) \approx 1 - \left(\frac{\lambda/\alpha}{1 + e^{\lambda(t-\mu)}} + \beta/\alpha \right), \quad \mu = \frac{1}{\lambda} \log \frac{\lambda^2}{\alpha} + \frac{1}{\lambda} \log \frac{s}{2v_R v_I}. \quad (19)$$

For lineage- i ($i = 2, 4$) let $v_R^{(i)}$ and $v_I^{(i)}$ be the mutation rates to rifampicin and isoniazid. As in [26], keeping all other parameters equal between the lineages we have the relative probability of resistance

$$\frac{\mathbb{P}(T^{(2)} \leq t)}{\mathbb{P}(T^{(4)} \leq t)} \approx \frac{v_R^{(2)} v_I^{(2)} (1 + 2v_R^{(4)} v_I^{(4)} e^{\lambda t} \alpha / \lambda^2 s)}{v_R^{(4)} v_I^{(4)} (1 + 2v_R^{(2)} v_I^{(2)} e^{\lambda t} \alpha / \lambda^2 s)}.$$

For times such that $v_R^{(i)} v_I^{(i)} e^{\lambda t} \ll 1$, $i = 2, 4$, we see the relative probability is simply $v_R^{(2)} v_I^{(2)} / v_R^{(4)} v_I^{(4)}$. Taking the mutation rates measured in [26], this gives the relative probability of 22.06. This agrees with the simulation results in [26] who reported an approximately 22 fold increase. The same problem was treated in [38] but for multidrug resistance to two drugs in a fixed order (i.e. resistance to rifampicin followed by isoniazid). Staying in the regime $v_R^{(i)} v_I^{(i)} e^{\lambda t} \ll 1$, if in (19), we replace the 2 by 1 (as only one path to resistance), and alter both λ^2/α by λ , and $e^{\lambda t}$ with M , our expression is the same to leading order as that given in [38] (equation 2) for the probability of resistance in a population of size M . This difference between fixed population size and fixed time results is the same as given in [16, 54].

Concluding remarks

Related work. Many previous studies have considered the probability of particular cell type emerging in a growing population, typically by a fixed time after the process starts or when the total population reaches a given size. The majority deal with the target population being one or two transitions away from the initial population (the root vertex in this paper) [16, 18, 23, 38, 55–58]. In particular we single out the pioneering work of Luria and Delbrück [55], which demonstrated the spontaneous nature of mutations by combining an appropriate mathematical model with bacterial experiments on phage resistance. The original model of [55], and its various incarnations [59, 60] have been extensively studied [16, 57, 61–65]. Its fully stochastic formulation, which is identical to our model for two vertices, is due to Bartlett [60]. This full model admits an explicit solution [57], and the model's asymptotic behavior has been recently explored [16, 64]. One of the simplest quantities of interest is the probability of no resistant cells at a fixed time, often called p_0 . It is a closely related quantity to our target hitting time described in (5), that is no resistant cell arises by a fixed time. Understanding p_0 provides a method to infer the per cell-rate at which resistance-conferring mutations are acquired, often termed the p_0 -method [66].

Some notable exceptions which consider the target population being greater than two transitions away are [25, 31, 54, 67–70] (ref. [31] is discussed above in the Imperfect drug penetration: combination therapy section). In [25], the same model as that presented here is numerically explored when all vertices have the same fitness ($\alpha(x) = \alpha, \beta(x) = \beta$ for all vertices

x), and the implications on multidrug therapy failure is emphasised. An efficient numerical method to compute the distribution of target hitting time and path probabilities via the iteration of generating functions is given in [68] with a focus on cancer initiation. Both [54, 67] are motivated by the accumulation of driver mutations in cancer, and so each transition leads to a fitness increase. For [67] the mean time of the k th driver mutation derived and compared to genetic data. The distribution of the time until the k th driver is sought in [54], whose methods are the closest to those used in this paper. There the authors employed an appealing approximation of the model studied in this paper in the path graph case (the approximation is that the seeding rate into vertex $x + 1$ from vertex x , which is exactly $v(x, x + 1)Z_x(t)$, is approximated by $v(x, x + 1)W_x^*e^{\lambda(x)t}$, for some sensibly chosen random variables W_x^*). Notably, again for small transition rates, the functional form of the distribution of the target hitting time is the same as that given in Theorem 1, however with a different median. The altered median derived in [54] demonstrates that when transitions lead to a fitness increase, the growth rates associated with the intermediate vertex populations have a far greater effect on the target hitting time when compared to the setting considered in this paper. The same type of approximate model was further used in [69] to investigate the target hitting time when transitions bring a fitness increase which is a random variable. Target hitting times on a path graph for a branching process with general offspring distribution, were also discussed in the very recent paper [70], but their main explicit results exclude cell death, and hold when successive transition rates along the path are assumed to be getting (infinitely) stronger.

The model and questions considered in this paper also arise frequently in the evolutionary emergence or evolutionary escape literature [71–75], with the notable distinction that in the evolutionary escape setting the root and intermediate vertex populations are destined to go extinct ($\lambda(x) < 0$ for $1 \leq x \leq N - 1$). This scenario is of interest when, for example, a homogeneous, pathogenic population (all residing at the root vertex in the language used in the present study) is treated with a therapy and must acquire sufficiently many mutations away from the root so as to become resistant (populate the target vertex). As in the setup of the present paper, the target hitting time is strongly controlled by the growth of the root population, $Z_1(t)$, which has positive growth rate, $\lambda > 0$, the target hitting times in the evolutionary escape setting are distinct from those given in the Time until target vertex is populated section. However for the escape probability (which is the probability of reaching the target), if there are multiple paths from the root to the target, the contribution of each path to the escape probability (termed the path value in [71, 72]) has an expression strikingly similar to the path weights discussed here for small transition rates (compare (2) with Eqs 6a–c in [72]). We might conjecture that for a specific path, the path value, as given in [72], is the unnormalised probability of reaching the target via the specified path, as we have demonstrated is the case with the path weights in Theorem 2 (this is implied in [72] Sec 2.5). Further connections surrounding the path distribution in these differing regimes is an interesting avenue for future work.

Summary. In this article we have considered a continuous time birth-death process with transitions. The time until a population of a particular target state arises was investigated. This target state is accessible via one or more transitions from the initial population. Which sequence of transitions leads to the target population arising was explored. This gave rise to a probability distribution over the paths composed of transitions between differing vertices. Motivated by applications, the setting when the initial population was most fit was focused on. The important factor in determining whether the target was reached via a specific path was seen to be the weight of the path. The weight of the path is composed of the transition rates and fitness costs associated with the intermediate vertices along the considered path.

Our primary contributions are the simple, explicit formulas for the target hitting time and the distribution of paths to the target, which are valid for small transition rates. When compared with previously given expressions for hitting times in branching processes, due to the reduction in complexity, we believe our results are easily interpretable and widely applicable. Further, to the best of our knowledge, the path distribution presented is the first analytic characterisation of evolutionary paths through a multitype branching process. The utility of our formulas was demonstrated on a variety of scenarios pertaining to the emergence of drug resistance. The biological relations revealed would have been difficult to deduce by a simulation based approach.

A shortcoming of our model is that the potential for unlimited growth and the lack of interactions between the cells is unrealistic. However in systems with large carrying capacities and abundant resources, our framework provides a convenient approximation and has been widely used. One possible route for assessing the importance of these neglected factors is to embed our questions in a similar setup to that used in the recent work [76], which included competition between cell types and a carrying capacity which the population could fluctuate around. The initial conditions in [76] were such that the model does not represent a growing population, so we do not compare their findings with the results of our Results section. Embedding our questions in a variant of the model of [76], such that the population is growing, might offer insight on the effect of these neglected aspects. Whenever these factors are deemed unimportant, or as a starting point for more complex models, this article provides quick and accessible results which may be used to guide and develop biological insight.

Materials and methods

Population growth

The main ingredient in proving Theorem 1, and consequently Theorem 2, is the long-time behaviour of the population of the initial and intermediate vertices. In this section we discuss this asymptotic behaviour.

Firstly we extend the definition of the total path weight given in the General framework section. Let the set of paths between the root and vertex x be denoted $\mathcal{P}_{1,x}$. Then using the definition of path weight (2), we let the total weight for vertex x be $\phi_x = \sum_{p \in \mathcal{P}_{1,x}} w(p)$ for $2 \leq x \leq N$. Further, let us denote the ratio of the total weight to fitness cost for vertex x as

$$\Phi_x = \frac{\phi_x}{\lambda - \lambda(x)}$$

for $2 \leq x \leq N - 1$. To allow us to conveniently describe the growth at vertex 1 also, we set $\Phi_1 = 1$.

The target population can only be founded by transitions from cells residing at the neighbouring vertices (vertices connected to the target by an edge). Therefore, understanding the population growth of cells at these vertices is needed to discuss the timing and manner in which cells at the target arise. At large times, this understanding is provided by the following theorem, concerning the population at the initial and intermediate vertices (it is useful to recall that $Z_x(t)$ is the number of cells at vertex x at time t and that we initiate with z cells at vertex 1). It follows from a more general result due to Janson [41], which we have tailored to our particular setting.

Theorem 3. *With probability one*

$$\lim_{t \rightarrow \infty} e^{-\lambda t} (Z_x(t))_{x=1}^{N-1} = W(\Phi_x)_{x=1}^{N-1}.$$

Here W is distributed as the sum of K independent exponential random variables with parameter λ/α , where K is binomial with z trials and success probability λ/α .

A short proof, demonstrating how Theorem 3 can be deduced from [41] is given in Section S2 of the [S1 Appendix](#) (see Theorem S1). The distribution of W may be called a binomial-Erlang mixture, being an Erlang distribution with a binomially distributed shape parameter. For the running example of a path graph on N vertices, we see that $\Phi_x = \prod_{i=2}^x \frac{v(i-1,i)}{\lambda - \lambda(i)}$ for $1 \leq x \leq N-1$. This is demonstrated in [Fig 3a](#). We note that in the case of $N=3$, Theorem 3 was previously given in [16] (Theorem 3.2).

Using Theorem 3 and that the immigration rate into the target vertex from any target neighbouring vertex x is $Z_x(t)v(x, N)$ we are able to prove the results of our Results section. Full proofs are provided in [S1 Appendix](#).

Supporting information

S1 Appendix. Mathematical details and proofs.
(PDF)

Acknowledgments

We thank David Cheek, Stefano Avanzini and the Edinburgh physics population dynamics group for helpful discussions and feedback.

Author Contributions

Conceptualization: Michael D. Nicholson, Tibor Antal.

Formal analysis: Michael D. Nicholson.

Investigation: Michael D. Nicholson.

Methodology: Michael D. Nicholson, Tibor Antal.

Supervision: Tibor Antal.

Writing – original draft: Michael D. Nicholson.

Writing – review & editing: Michael D. Nicholson, Tibor Antal.

References

1. J. Berestycki, E. Brunet, and Z. Shi, "The number of accessible paths in the hypercube," *Bernoulli*, vol. 22, pp. 653–680, 05 2016. <https://doi.org/10.3150/14-BEJ641>
2. Beerenwinkel N., Antal T., Dingli D., Traulsen A., Kinzler K. W., Velculescu V. E., Vogelstein B., and Nowak M. A., "Genetic progression and the waiting time to cancer," *PLOS Computational Biology*, vol. 3, pp. 1–8, 11 2007. <https://doi.org/10.1371/journal.pcbi.0030225>
3. Weinreich D. M., Delaney N. F., DePristo M. A., and Hartl D. L., "Darwinian evolution can follow only very few mutational paths to fitter proteins," *Science*, vol. 312, no. 5770, pp. 111–114, 2006. <https://doi.org/10.1126/science.1123539>
4. Weissman D. B., Desai M. M., Fisher D. S., and Feldman M. W., "The rate at which asexual populations cross fitness valleys," *Theoretical Population Biology*, vol. 75, no. 4, pp. 286–300, 2009. Sam Karlin: Special Issue. <https://doi.org/10.1016/j.tpb.2009.02.006>
5. Durrett R., Schmidt D., and Schweinsberg J., "A waiting time problem arising from the study of multi-stage carcinogenesis," *The Annals of Applied Probability*, vol. 19, no. 2, pp. 676–718, 2009. <https://doi.org/10.1214/08-AAP559>
6. Gillespie J. H., "A simple stochastic gene substitution model," *Theoretical Population Biology*, vol. 23, no. 2, pp. 202–215, 1983. [https://doi.org/10.1016/0040-5809\(83\)90014-X](https://doi.org/10.1016/0040-5809(83)90014-X)

7. Zagorski M., Burda Z., and Waclaw B., "Beyond the hypercube: Evolutionary accessibility of fitness landscapes with realistic mutational networks," *PLOS Computational Biology*, vol. 12, pp. 1–18, 12 2016. <https://doi.org/10.1371/journal.pcbi.1005218>
8. Gokhale C. S., Iwasa Y., Nowak M. A., and Traulsen A., "The pace of evolution across fitness valleys," *Journal of Theoretical Biology*, vol. 259, no. 3, pp. 613–620, 2009. <https://doi.org/10.1016/j.jtbi.2009.04.011>
9. de Visser J. A. G. M. and Krug J., "Empirical fitness landscapes and the predictability of evolution," *Nature Reviews Genetics*, vol. 15, p. 480, June 2014. <https://doi.org/10.1038/nrg3744>
10. Tadrowski A. C., Evans M. R., and Waclaw B., "Phenotypic Switching Can Speed up Microbial Evolution," *Scientific Reports*, vol. 8, no. 1, p. 8941, 2018. <https://doi.org/10.1038/s41598-018-27095-9>
11. Hermesen R. and Hwa T., "Sources and sinks: A stochastic model of evolution in heterogeneous environments," *Phys. Rev. Lett.*, vol. 105, p. 248104, Dec 2010. <https://doi.org/10.1103/PhysRevLett.105.248104>
12. Nowak M. A., Komarova N. L., Sengupta A., Jallepalli P. V., Shih I.-M., Vogelstein B., and Lengauer C., "The role of chromosomal instability in tumor initiation," *Proceedings of the National Academy of Sciences*, vol. 99, no. 25, pp. 16226–16231, 2002. <https://doi.org/10.1073/pnas.202617399>
13. Nichol D., Jeavons P., Fletcher A. G., Bonomo R. A., Maini P. K., Paul J. L., Gatenby R. A., Anderson A. R., and Scott J. G., "Steering evolution with sequential therapy to prevent the emergence of bacterial antibiotic resistance," *PLOS Computational Biology*, vol. 11, pp. 1–19, 09 2015. <https://doi.org/10.1371/journal.pcbi.1004493>
14. Williams M. J., Werner B., Barnes C. P., Graham T. A., and Sottoriva A., "Identification of neutral tumor evolution across cancer types," *Nature Genetics*, vol. 48, pp. 238–244, 2016. <https://doi.org/10.1038/ng.3489>
15. Bozic I., Gerold J. M., and Nowak M. A., "Quantifying clonal and subclonal passenger mutations in cancer evolution," *PLOS Computational Biology*, vol. 12, pp. 1–19, 02 2016. <https://doi.org/10.1371/journal.pcbi.1004731>
16. Cheek D. and Antal T., "Mutation frequencies in a birth–death branching process," *Ann. Appl. Probab.*, vol. 28, pp. 3922–3947, 12 2018. <https://doi.org/10.1214/18-AAP1413>
17. Gao R., Davis A., McDonald T. O., Sei E., Shi X., Wang Y., Tsai P.-C., Casasent A., Waters J., Zhang H., Meric-Bernstam F., Michor F., and Navin N. E., "Punctuated copy number evolution and clonal stasis in triple-negative breast cancer," *Nature Genetics*, vol. 48, p. 1119, Aug. 2016. <https://doi.org/10.1038/ng.3641>
18. Haeno H., Iwasa Y., and Michor F., "The evolution of two mutations during clonal expansion," *Genetics*, vol. 177, no. 4, pp. 2209–2221, 2007. <https://doi.org/10.1534/genetics.107.078915>
19. Haeno H., Gonen M., Davis M., Herman J., Iacobuzio-Donahue C., and Michor F., "Computational modeling of pancreatic cancer reveals kinetics of metastasis suggesting optimum treatment strategies," *Cell*, vol. 148, no. 1, pp. 362–375, 2012. <https://doi.org/10.1016/j.cell.2011.11.060>
20. Nicholson M. D. and Antal T., "Universal asymptotic clone size distribution for general population growth," *Bull Math Biol*, vol. 78, pp. 2243–2276, 2016. <https://doi.org/10.1007/s11538-016-0221-x>
21. Dewanji A., Jeon J., Mexa R., and Luebeck E. G., "Number and size distribution of colorectal adenomas under the multistage clonal expansion model of cancer," *PLoS Computational Biology*, vol. 7, no. 10, 2011. <https://doi.org/10.1371/journal.pcbi.1002213>
22. Reiter J. G., Makohon-Moore A. P., Gerold J. M., Heyde A., Attiyeh M. A., Kohutek Z. A., Tokheim C. J., Brown A., DeBlasio R. M., Niyazov J., Zucker A., Karchin R., Kinzler K. W., Iacobuzio-Donahue C. A., Vogelstein B., and Nowak M. A., "Minimal functional driver gene heterogeneity among untreated metastases," *Science*, vol. 361, no. 6406, pp. 1033–1037, 2018. <https://doi.org/10.1126/science.aat7171>
23. Bozic I., Reiter J. G., Allen B., Antal T., Chatterjee K., Shah P., Moon Y. S., Yaqubie A., Kelly N., Le D. T., Lipson E. J., Chapman P. B., Diaz J., Luis A., Vogelstein B., and Nowak M. A., "Evolutionary dynamics of cancer in response to targeted combination therapy," *eLife*, vol. 2, p. e00747, jun 2013. <https://doi.org/10.7554/eLife.00747>
24. Kessler D. A., Austin R. H., and Levine H., "Resistance to chemotherapy: patient variability and cellular heterogeneity," *Cancer research*, vol. 74, no. 17, pp. 4663–70, 2014. <https://doi.org/10.1158/0008-5472.CAN-14-0118>
25. Komarova N. L. and Wodarz D., "Drug resistance in cancer: Principles of emergence and prevention," *Proceedings of the National Academy of Sciences*, vol. 102, no. 27, pp. 9714–9719, 2005. <https://doi.org/10.1073/pnas.0501870102>
26. Ford C. B., Shah R. R., Maeda M. K., Gagneux S., Murray M. B., Cohen T., Johnston J. C., Gardy J., Lipsitch M., and Fortune S. M., "Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis," *Nature Genetics*, vol. 45, no. 7, pp. 784–790, 2013. <https://doi.org/10.1038/ng.2656>

27. Wilkinson R. D. and Tavaré S., "Estimating primate divergence times by using conditioned birth-and-death processes," *Theoretical Population Biology*, vol. 75, no. 4, pp. 278–285, 2009. Sam Karlin: Special Issue. <https://doi.org/10.1016/j.tpb.2009.02.003>
28. Greulich P., Waclaw B., and Allen R. J., "Mutational pathway determines whether drug gradients accelerate evolution of drug-resistant cells," *Phys. Rev. Lett.*, vol. 109, p. 088101, Aug 2012. <https://doi.org/10.1103/PhysRevLett.109.088101>
29. Hermesen R., Deris J. B., and Hwa T., "On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient," *Proceedings of the National Academy of Sciences*, vol. 109, no. 27, pp. 10775–10780, 2012. <https://doi.org/10.1073/pnas.1117716109>
30. Fu F., Nowak M. A., and Bonhoeffer S., "Spatial Heterogeneity in Drug Concentrations Can Facilitate the Emergence of Resistance to Cancer Therapy," *PLoS Computational Biology*, vol. 11, no. 3, 2015. <https://doi.org/10.1371/journal.pcbi.1004142>
31. Moreno-Gamez S., Hill A. L., Rosenbloom D. I. S., Petrov D. A., Nowak M. A., and Pennings P. S., "Imperfect drug penetration leads to spatial monotherapy and rapid evolution of multidrug resistance," *Proceedings of the National Academy of Sciences*, vol. 112, no. 22, pp. E2874–E2883, 2015. <https://doi.org/10.1073/pnas.1424184112>
32. Bozic I. and Nowak M. A., "Resisting resistance," *Annual Review of Cancer Biology*, vol. 1, no. 1, pp. 203–221, 2017. <https://doi.org/10.1146/annurev-cancerbio-042716-094839>
33. Andersson D. I. and Hughes D., "Antibiotic resistance and its cost: is it possible to reverse resistance?," *Nature Reviews Microbiology*, 2010. <https://doi.org/10.1038/nrmicro2319>
34. Hughes D. and Andersson D. I., "Evolutionary consequences of drug resistance: Shared principles across diverse targets and organisms," *Nature Reviews Genetics*, vol. 16, no. 8, pp. 459–471, 2015. <https://doi.org/10.1038/nrg3922>
35. Enriquez-Navas P. M., Kam Y., Das T., Hassan S., Silva A., Foroutan P., Ruiz E., Martinez G., Minton S., Gillies R. J., and Gatenby R. A., "Exploiting evolutionary principles to prolong tumor control in pre-clinical models of breast cancer," *Science Translational Medicine*, vol. 8, no. 327, pp. 327ra24–327ra24, 2016. <https://doi.org/10.1126/scitranslmed.aad7842>
36. Tomasetti C., Vogelstein B., and Parmigiani G., "Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation," *Proceedings of the National Academy of Sciences*, 2013. <https://doi.org/10.1073/pnas.1221068110>
37. Lee H., Popodi E., Tang H., and Foster P. L., "Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing," *Proceedings of the National Academy of Sciences*, vol. 109, no. 41, pp. E2774–E2783, 2012. <https://doi.org/10.1073/pnas.1210309109>
38. Colijn C., Cohen T., Ganesh A., and Murray M., "Spontaneous emergence of multiple drug resistance in tuberculosis before and during therapy," *PLOS ONE*, vol. 6, pp. 1–7, 03 2011. <https://doi.org/10.1371/journal.pone.0018327>
39. Athreya K. B. and Ney P. E., *Branching Processes*. Dover Publications, 2004.
40. Pearson J. E., Krapivsky P., and Perelson A. S., "Stochastic theory of early viral infection: Continuous versus burst production of virions," *PLOS Computational Biology*, vol. 7, pp. 1–17, 02 2011. <https://doi.org/10.1371/journal.pcbi.1001058>
41. Janson S., "Functional limit theorems for multitype branching processes and generalized Pólya urns," *Stochastic Processes and their Applications*, vol. 110, no. 2, pp. 177–245, 2004. <https://doi.org/10.1016/j.spa.2003.12.002>
42. Deresinski S., Vancomycin in Combination with Other Antibiotics for the Treatment of Serious Methicillin-Resistant *Staphylococcus aureus* Infections, *Clinical Infectious Diseases*, 49 (2009), pp. 1072–1079. <https://doi.org/10.1086/605572> PMID: 19725789
43. Warner D. F. and Mizrahi V., "Tuberculosis chemotherapy: The influence of bacillary stress and damage response pathways on drug efficacy," 2006. <https://doi.org/10.1128/CMR.00060-05>
44. Primeau A. J., Rendon A., Hedley D., Lilje L., and Tannock I. F., "The distribution of the anticancer drug doxorubicin in relation to blood vessels in solid tumors," *Clinical Cancer Research*, vol. 11, no. 24, pp. 8782–8788, 2005. <https://doi.org/10.1158/1078-0432.CCR-05-1664>
45. Cihlar T. and Fordyce M., "Current status and prospects of hiv treatment," *Current Opinion in Virology*, vol. 18, pp. 50–56, 2016. Antiviral strategies • Virus structure and expression. <https://doi.org/10.1016/j.coviro.2016.03.004>
46. Tamma P. D., Cosgrove S. E., and Maragakis L. L., "Combination therapy for treatment of infections with gram-negative bacteria," *Clinical Microbiology Reviews*, vol. 25, no. 3, pp. 450–470, 2012. <https://doi.org/10.1128/CMR.05041-11>

47. Leder K., Foo J., Skaggs B., Gorre M., Sawyers C. L., and Michor F., "Fitness conferred by bcr-abl kinase domain mutations determines the risk of pre-existing resistance in chronic myeloid leukemia," *PLOS ONE*, vol. 6, pp. 1–11, 11 2011. <https://doi.org/10.1371/journal.pone.0027682>
48. Tlsty T. D., Margolin B. H., and Lum K., "Differences in the rates of gene amplification in nontumorigenic and tumorigenic cell lines as measured by Luria-Delbruck fluctuation analysis," *Proc Natl Acad Sci U S A*, vol. 86, no. 23, pp. 9441–9445, 1989. <https://doi.org/10.1073/pnas.86.23.9441>
49. Michor F., Hughes T. P., Iwasa Y., Branford S., Shah N. P., Sawyers C. L., and Nowak M. A., "Dynamics of chronic myeloid leukaemia," *Nature*, vol. 435, no. 7046, pp. 1267–1270, 2005. <https://doi.org/10.1038/nature03669>
50. Bixby D. and Talpaz M., "Seeking the causes and solutions to imatinib-resistance in chronic myeloid leukemia," *Leukemia*, vol. 25, p. 7, Nov. 2010. <https://doi.org/10.1038/leu.2010.238>
51. Tipping A. J., Manon F. X., Lagarde V., Goldman J. M., and Melo J. V., "Restoration of sensitivity to STI571 in STI571-resistant chronic myeloid leukemia cells," *Blood*, vol. 98, no. 13, pp. 3864–3867, 2001. <https://doi.org/10.1182/blood.V98.13.3864>
52. Hall A. R., Iles J. C., and MacLean R. C., "The fitness cost of rifampicin resistance in *Pseudomonas aeruginosa* depends on demand for RNA polymerase," *Genetics*, vol. 187, no. 3, pp. 817–822, 2011. <https://doi.org/10.1534/genetics.110.124628>
53. Ward H., Perron G. G., and MacLean R. C., "The cost of multiple drug resistance in *Pseudomonas aeruginosa*," *Journal of Evolutionary Biology*, vol. 22, no. 5, pp. 997–1003, 2009.
54. Durrett R. and Moseley S., "Evolution of resistance and progression to disease during clonal expansion of cancer," *Theoretical Population Biology*, vol. 77, no. 1, pp. 42–48, 2010. <https://doi.org/10.1016/j.tpb.2009.10.008>
55. Luria S. E. and Delbrück M., "Mutations of bacteria from virus sensitivity to virus resistance," *Genetics*, vol. 48, no. 6, pp. 491–511, 1943.
56. Iwasa Y., Nowak M. A., and Michor F., "Evolution of resistance during clonal expansion," *Genetics*, vol. 172, no. 4, pp. 2557–2566, 2006. <https://doi.org/10.1534/genetics.105.049791>
57. Antal T. and Krapivsky P. L., "Exact solution of a two-type branching process: models of tumor progression," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2011, no. 08, p. P08018, 2011. <https://doi.org/10.1088/1742-5468/2011/08/P08018>
58. Denes J. and Krewski D., "An exact representation for the generating function for the moogavkar-venzon-kudson two-stage model of carcinogenesis with stochastic stem cell growth," *Mathematical Biosciences*, vol. 131, no. 2, pp. 185–204, 1996. [https://doi.org/10.1016/0025-5564\(95\)00046-1](https://doi.org/10.1016/0025-5564(95)00046-1)
59. Lea D. E. and Coulson C. A., "The distribution of the numbers of mutants in bacterial populations," *Journal of Genetics*, vol. 49, no. 3, pp. 264–285, 1949. <https://doi.org/10.1007/BF02986080>
60. Bartlett M., *An Introduction to Stochastic Processes*. Cambridge University Press, 1955.
61. Mandelbrot B., "A Population Birth-and-Mutation Process, I: Explicit Distributions for the Number of Mutants in an Old Culture of Bacteria," *Journal of Applied Probability*, vol. 11, no. 3, pp. pp. 437–444, 1974. <https://doi.org/10.2307/3212688>
62. Möhle M., "Convergence Results for Compound Poisson Distributions and Applications to the Standard Luria-Delbrück Distribution," *Journal of Applied Probability*, vol. 42, no. 3, pp. pp. 620–631, 2005. <https://doi.org/10.1239/jap/1127322016>
63. Ma W. T., Sandri G. V., and Sarkar S., "Analysis of the Luria-Delbruck Distribution using discrete convolution powers," *Journal of Applied Probability*, vol. 29, no. 2, pp. 255–267, 1992. <https://doi.org/10.2307/3214564>
64. Kessler D. A. and Levine H., "Large population solution of the stochastic Luria-Delbruck evolution model," *Proceedings of the national Academy of Science USA*, vol. 110, no. 29, pp. 11682–11687, 2013. <https://doi.org/10.1073/pnas.1309667110>
65. Keller P. and Antal T., "Mutant number distribution in an exponentially growing population," *J. Stat. Mech.* P01011, no. 1, 2015.
66. Rosche W. A. and Foster P. L., "Determining mutation rates in bacterial populations," *Methods*, vol. 20, pp. 4–17, 2000. <https://doi.org/10.1006/meth.1999.0901>
67. Bozic I., Antal T., Ohtsuki H., Carter H., Kim D., Chen S., Karchin R., Kinzler K. W., Vogelstein B., and Nowak M. A., "Accumulation of driver and passenger mutations during tumor progression," *Proceedings of the National Academy of Sciences*, vol. 107, no. 43, pp. 18545–18550, 2010. <https://doi.org/10.1073/pnas.1010978107>
68. Bauer B. and Gokhale C. S., "Repeatability of evolution on epistatic landscapes," *Scientific Reports*, vol. 5, pp. 9607 EP –, 05 2015. <https://doi.org/10.1038/srep09607>

69. Durrett R., Foo J., Leder K., Mayberry J., and Michor F., "Intratumor heterogeneity in evolutionary models of tumor progression," *Genetics*, vol. 188, no. 2, pp. 461–477, 2011. <https://doi.org/10.1534/genetics.110.125724>
70. Chaumont L. and Nguyen T. N. A., "On mutations in the branching model for multitype populations," *Advances in Applied Probability*, vol. 50, no. 2, pp. 543–564, 2018. <https://doi.org/10.1017/apr.2018.24>
71. Iwasa Y., Michor F., and Nowak M. A., "Evolutionary dynamics of escape from biomedical intervention," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, vol. 270, no. 1533, pp. 2573–2578, 2003. <https://doi.org/10.1098/rspb.2003.2539>
72. Iwasa Y., Michor F., and Nowak M. A., "Evolutionary dynamics of invasion and escape," *Journal of Theoretical Biology*, vol. 226, no. 2, pp. 205–214, 2004. <https://doi.org/10.1016/j.jtbi.2003.08.014>
73. Serra M. C. and Haccou P., "Dynamics of escape mutants," *Theoretical Population Biology*, vol. 72, no. 1, pp. 167–178, 2007. <https://doi.org/10.1016/j.tpb.2007.01.005>
74. Sagitov S. and Conceição Serra M., "Multitype bienaymé–galton–watson processes escaping extinction," *Advances in Applied Probability*, vol. 41, 03 2009.
75. Alexander H. K., "Conditional distributions and waiting times in multitype branching processes," *Advances in Applied Probability*, vol. 45, no. 3, pp. 692–718, 2013. <https://doi.org/10.1239/aap/1377868535>
76. A. Bovier, L. Coquille, and C. Smadi, "Crossing a fitness valley as a metastable transition in a stochastic population model," *arXiv e-prints*, p. arXiv:1801.06473, Jan. 2018.